



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 959 128 A1**

(12) **EUROPEAN PATENT APPLICATION**  
published in accordance with Art. 158(3) EPC

(43) Date of publication:  
24.11.1999 Bulletin 1999/47

(21) Application number: 97933003.2

(22) Date of filing: 24.07.1997

(51) Int. Cl.<sup>6</sup>: C12N 9/42, C12N 15/56

(86) International application number:  
PCT/JP97/02561

(87) International publication number:  
WO 98/03640 (29.01.1998 Gazette 1998/04)

(84) Designated Contracting States:  
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE

(30) Priority: 24.07.1996 JP 19497496

(71) Applicant:  
MEIJI SEIKA KABUSHIKI KAISHA  
Tokyo 104 (JP)

(72) Inventors:  
• MURASHIMA, Kouichirou,  
Bioscience Labs.  
Sakado-shi, Saitama 350-02 (JP)  
• MORIYA, Tatsuki,  
Pharmaceutical Technology Labs.  
Odawara-shi, Kanagawa 250-01 (JP)  
• HAMAYA, Toru,  
Bioscience Labs.  
Sakado-shi, Saitama 350-02 (JP)

• KOGA, Jinichiro,  
Bioscience Labs.  
Sakado-shi, Saitama 350-02 (JP)  
• SUMIDA, Naomi,  
Pharmaceutical Technology Labs.  
Odawara-shi, Kanagawa 250-01 (JP)  
• AOYAGI, Kaoru,  
Pharmaceutical Technology Labs.  
Odawara-shi, Kanagawa 250-01 (JP)  
• MURAKAMI, Takeshi,  
Pharmaceutical Technology Labs.  
Odawara-shi, Kanagawa 250-01 (JP)  
• KONO, Toshiaki,  
Bioscience Labs.  
Sakado-shi, Saitama 350-02 (JP)

(74) Representative:  
Hall, Marina et al  
Elkington and Fife  
Prospect House,  
8 Pembroke Road  
Sevenoaks, Kent TN13 1XR (GB)

(54) **ENZYME ENDOGLUCANASE AND CELLULASE PREPARATIONS CONTAINING THE SAME**

(57) A highly active cellulase which is to be preferably employed in eliminating fluffs from cellulose-containing fibers, denier reduction and bleaching denim-dyed cellulose fibers; and a gene thereof. This novel cellulase NCE4 isolated from *Humicola insolens* is a highly active one usable in treating various cellulose-containing fibers.

EP 0 959 128 A1

**Description****[Background of the Invention]**5 **Field of the Invention**

**[0001]** The present invention relates to cellulase and a cellulase preparation containing the same, as well as a removal of nap of cellulose-containing fibers, a process for reducing cellulose-containing fibers and a process for decoloring denim-dyed cellulose-containing fibers by using the cellulase or cellulase preparation.

10

**Background Art**

**[0002]** Cellulose-containing fibers are treated with cellulase to give desired properties to those fibers. For example, in the fiber industry, treatment is performed using cellulase in order to improve the texture and appearance of cellulose-containing fibers, and to give the appearance of being stonewashed to denim-dyed cellulose-containing fibers.

15

**[0003]** In addition, Tencel, a reproduced cellulose fiber that is spun wood pulp-based cellulose dissolved in an organic solvent, has attracted attention in recent years due to properties such as high strength, water absorbing capacity and the like, and its manufacturing method which causes very little environmental contamination. However, since nap is produced during the manufacturing process of Tencel, Tencel has low product value as a fiber if left as is. Therefore, methods for removing nap formed during the manufacturing process by using cellulase have been proposed.

20

**[0004]** At present, cellulases primarily originating in *Trichoderma* species and *Humicola* species of wood-rotting fungi are used for treatment of cellulose-containing fibers. However, a large amount of cellulase is required under the present conditions to demonstrate the desired effect on the fibers.

25

**[0005]** It is likely that the cost of this processing would be able to be reduced, if it were possible to be able to improve the texture and appearance of cellulose-containing fibers, give the appearance of being stone-washed to denim-dyed cellulose-containing fibers, and remove nap from Tencel by using highly active cellulase in smaller amount than the conventional amount.

30

**[0006]** Furthermore, an example of a cellulase originating in *Humicola*, it is shown in WO 91/17243 (Japanese Patent Laid-Open Publication No. 5-509223) that the endoglucanase gene of 43 kD was isolated from *Humicola insolens* DSM1800 strain, and that its base sequence was determined.

**[Summary of the Invention]**

**[0007]** The inventors of the present invention have now isolated from *Humicola insolens* a novel highly active cellulase and its gene, which can be used extremely advantageously in the treatment of various cellulose-containing fibers. The present invention is based on those findings.

35

**[0008]** Thus, an object of the present invention is to provide a novel cellulase and its gene.

40

**[0009]** Moreover, another object of the present invention is to provide a process for removing nap of cellulose-containing fibers using a novel cellulase, a process for reducing cellulose-containing fibers and a process for decoloring denim-dyed, cellulose-containing fibers.

45

**[0010]** The novel cellulase according to the present invention is a protein or a modified protein having a portion of an amino acid sequence shown in SEQ ID No. 1 or a sequence from position 1 to position 284 of the amino acid sequence shown in SEQ ID No. 1.

50

**[0011]** In addition, according to the present invention, there is provided a process for removing nap of cellulose-containing fibers, a process for reducing cellulose-containing fibers, and a process for decoloring denim-dyed, cellulose-containing fibers comprising the step of contacting the protein or modified protein of the present invention With the denim-dyed, cellulose-containing fibers.

**[Detailed Description of the Invention]**

55

**Microorganism Deposition**

**[0012]** *E. coli* JM109 strain transformed by plasmid pNCE4Sal (see Example A5) containing the cellulase gene according to the present invention has been deposited in the National Institute of Bioscience and Human-Technology, Ministry of International Trade and Industry (1-1-3 Higashi, Tsukuba-shi, Ibaraki-ken, Japan) under accession No. FERM BP-5976 (original deposit: FERM P-15732, date of original deposit: July 12, 1996).

## Cellulase and its Gene

[0013] The cellulose enzyme according to the present invention is a protein having a portion or a sequence from position 1 to position 284 of the sequence shown in SEQ ID No. 1. In the present invention, a portion of the sequence shown in SEQ ID No. 1 refers to a partial sequence, for example, that has a length available for a probe, or that still maintains cellulase activity, particularly endoglucanase activity.

[0014] In addition, the present invention also includes a protein having a portion or all of the amino acid sequence from position -22 to position -1 of SEQ ID No. 1 on the N-terminal of the above protein. The amino acid sequence from position -22 to position -1 of SEQ ID No. 1 is considered to be a signal peptide. Therefore, in addition to the partial sequence that retains signal peptide activity, a portion of the sequence includes the sequence that remains on the N-terminal as a result of the processing in which the processing position may vary depending on the type of a host.

[0015] Moreover, modified proteins of the above protein are also encompassed by the scope of the present invention. In the present invention, a modified protein refers to a protein in which a modification, such as the addition, insertion, deletion or substitution of one or a plurality of amino acids, has occurred in the amino acid sequence of the protein, while still retaining cellulase activity, particularly endoglucanase activity.

[0016] The sequence from position 1 to position 284 of the amino acid sequence shown in SEQ ID No. 1 is hereinafter referred to as cellulase NCE4, and its gene is referred to as the cellulase NCE4 gene.

[0017] The cellulase according to the present invention has high activity, and thus allows for the desired effects in various uses even in small amounts. For example, according to a preferred embodiment of the present invention, in comparison with crude cellulase prepared from a culture liquid of *Humicola insolens*, NCE4 shows equivalent nap removal effects on cellulose fibers at a protein mass of about 1/100 that of the crude cellulase, equivalent denim-dyed cellulose-containing fiber decoloring effects at a protein mass about 1/25, and equivalent cellulose fiber reduction effects at a protein mass of about 1/5. Accordingly, it can be used for more efficient and economical treatment of cellulose fibers.

[0018] According to another embodiment of the present invention, there is provided a DNA sequence that codes for the amino acid sequence of the above protein. A typical sequence of the DNA sequence has a portion or all of the base sequence shown in SEQ ID No. 2.

[0019] The base sequence shown in SEQ ID No. 2 has an open reading frame starting at ATG at positions 118 to 120 and terminating at TAA at positions 1089 to 1091. In addition, the base sequence from position 171 to position 173 corresponds to the mature protein composed of 284 residues. Moreover, the presence of an intron has been confirmed in the base sequence of SEQ ID No. 2 (see Example A7).

[0020] When the amino acid sequence of a protein is given, the DNA sequence coding for that sequence is easily defined, and various base sequences that code for a portion or all of the amino acid sequence shown in SEQ ID No. 1 can be selected. Thus, a DNA sequence coding for a portion or all of the amino acid sequence shown in SEQ ID No. 1 according to the present invention refers to a sequence having any degenerate codons coding for the same amino acids in a base sequence, in addition to a portion or all of the base sequence shown in SEQ ID No. 2.

[0021] The DNA according to the present invention may be of natural origin or be completely synthesized. In addition, it may also be synthesized using a portion of DNA of natural origin. Typical methods for obtaining DNA include a method of obtaining the DNA from a *Humicola insolens* chromosome library or cDNA library routinely employed in the field of genetic engineering, an example of which includes screening using a suitable DNA probe prepared based on information about a partial amino acid sequence. In addition, the DNA can also be obtained from the deposited microorganism described above.

Expression Vector and Transformed Microorganism

[0022] According to the present invention, there is provided an expression vector that contains the DNA sequence according to the present invention in a state that allows the DNA sequence to replicate within a host microorganism and to express a protein encoded by the DNA sequence. Moreover, according to the present invention, there is provided a microorganism that is transformed by the expression vector. There are no particular limitations on this host-vector system, and for example, systems using *E. coli*, Actinomycetes, yeast or molds, as well as fused protein expression systems with other proteins using these microorganisms, can be used. Methods routinely employed in the field of genetic engineering can be used for the procedure and method for constructing the vector according to the present invention.

[0023] The expression vector of the present invention may contain, in addition to the DNA sequence of the present invention, a DNA sequence that controls the expression of a desired protein or a selectable marker to ensure the introduction of the vector into a host and the expression of the desired protein. Since the base sequence of SEQ ID No. 2 is considered to contain there regulatory sequence, the use of this base sequence may be advantageous. In addition, the expression vector of the present invention may contain DNA sequences that code for cellulase in a repetitive form (tandem). The expression vector of the present invention that contains the DNA sequence of the present invention may be constructed in accordance with routine methods. In addition, transformation of a microorganism by the expression vec-

tor of the present invention may also be carried out in accordance with methods routinely employed in the art.

[0024] The transformant of the present invention can be cultured in a suitable medium. The protein according to the present invention can be isolated from this culture. Thus, according to another embodiment of the present invention, there is provided a method for preparing the novel protein of the present invention. The culturing of the transformants and its conditions may be basically similar to those of the microorganism used. In addition, recovery and purification of the novel protein according to the present invention from the culture can be performed in accordance with routine methods.

#### Cellulase Applications/Cellulase Preparation

[0025] The cellulase according to the present invention is used in various applications in its original form or as a cellulase preparation. In particular, it is used to give desired properties to cellulose fibers. More specifically, it is used for removing nap of cellulose-containing fibers, performing weight reduction processing, and performing decoloring processing of denim-dyed, cellulose-containing fibers.

[0026] A cellulase preparation may be prepared by mixing the cellulase according to the present invention with ingredients generally contained in cellulase preparations, examples of which include excipients (e.g., lactose, sodium chloride, sorbitol), surfactants and antiseptics.

[0027] A process for removing nap of cellulose-containing fibers, a process for reducing cellulose-containing fibers, and a process for decoloring denim-dyed, cellulose-containing fibers using the cellulase or cellulase preparation according to the present invention can be carried out by contacting the cellulase of the present invention with cellulose-containing fibers.

[0028] Although contact temperature, the amount of cellulase and other conditions may be decided depending on various conditions in a system, for example, in nap removal of cellulose-containing fibers, treatment can be performed at about 50-60°C using cellulase at a protein concentration of 5-50 mg/liter. In the process for reducing cellulose-containing fibers, treatment can be performed at about 50-60°C using cellulase at a protein concentration of 100-300 mg/liter. Moreover, in the process for decoloring denim-dyed, cellulose-containing fibers, treatment can be performed at about 50-60°C using cellulase at a protein concentration of 2-10 mg/liter.

#### [Examples]

[0029] The present invention will be further explained in view of the following examples that are not meant to limit the invention.

#### Example A1: Isolation and Purification of an Ingredient Having the Tencel Nap Removing Activity from *Humicola insolens*

[0030] *Humicola insolens* MN200-1 was cultured at 37°C in (N) medium (5.0% Avicel, 2.0% yeast extract, 0.1% polypeptone, 0.03% calcium chloride, 0.03% magnesium sulfate, pH 6.8). After culturing for 7 days, by centrifuging the resulting culture liquid at 7000 rpm for 20 minutes, the microorganisms were removed and the culture supernatant was used as the crude cellulase preparation.

[0031] This crude cellulase preparation was applied to hydrophobic chromatography (Phenyl-Sepharose High-Performance 16/100, Pharmacia-Biotech), and fractionated by eluting with aqueous ammonium sulfate over a concentration gradient of 1 to 0 M in 50 mM phosphate buffer (pH 7.0). Since potent Tencel nap removing activity was observed in the fraction obtained during a concentration gradient of 0.1 to 0 M, that fraction was again applied to hydrophobic chromatography (Phenyl-Sepharose High-Performance 16/100) and eluted with aqueous ammonium sulfate over a concentration gradient of 0.4 to 0 M in 50 mM phosphate buffer (pH 7.0) followed by removal of the active fraction.

[0032] Then, this fraction was applied to reversed-phase chromatography (Source 15 ISO, Pharmacia-Biotech) and fractionated by eluting with aqueous ammonium sulfate over a concentration gradient of 1 to 0 M in 50 mM phosphate buffer (pH 7.0). Since the fraction obtained during a concentration of 0 M was observed to exhibit potent Tencel nap removing activity, that fraction was further applied to reversed-phase chromatography (Source 15 PHE, Pharmacia-Biotech), eluted with aqueous ammonium sulfate over a concentration gradient of 1 to 0 M in 50 mM phosphate buffer (pH 7.0), and the fraction exhibiting potent Tencel nap removing activity was isolated as purified enzyme NCE4. This NCE4 exhibited a single band at a molecular weight of 43 kDa in SDS-PAGE.

## Example A2: Partial Amino Acid Sequence of Cellulase NCE4

## (1) Identification of N-terminal amino acid residues

5 [0033] Column chromatography (column: RESOURCE (Trade name) RPC 3 ml. 5-60% acetonitrile gradient containing 0.1% TFA) was performed with an FPLC system (Pharmacia-Biotech) followed by removal of the main peak to determine the N-terminal amino acid sequence of the protein purified in Example 1.

[0034] After lyophilizing the above fraction, it was dissolved in a small amount of water followed by electrophoresis using 8% Gel SDS-PAGE Mini (TEFCO). After electrically transferring the protein to PVDF membrane (Millipore) using a Multiphore II electrophoresis system (Pharmacia-Biotech), and staining with Coomassie brilliant blue R-250 (Nakalai tesque) the membrane was decolorized, washed with water and air dried. The portion at which the protein having a molecular weight of 43 kDa was blotted was cut out and applied to the Model 492 Protein Sequencer (Perkin-Elmer) to determine 15 residues of the N-terminal amino acid sequence. The resulting sequence was as shown below.

15 N-terminal amino acid sequence: Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp-(Cys)-(Cys)-Lys-Pro-Ser (15 residues)

## (2) Peptide Mapping

[0035] After lyophilizing the protein purified by FPLC in (1) above, it was dissolved in 100 mM ammonium bicarbonate buffer (pH 8.0). Approximately 1/20 mole volume of trypsin (Promega) relative to the amount of protein was added and allowed to react for 48 hours at 37°C. Column chromatography (column: C8 220 x 2.1 mm, 0.1% TFA, 0% acetonitrile to 0.085% TFA, 35% acetonitrile gradient) was performed on this dissolution product with the Model 172µ Preparative HPLC System (Perkin-Elmer) to obtain three types of peptides. The amino acid sequences of the resulting peptide fragments were determined using the above protein sequencer. Those results were as shown below.

25

TP-1: Tyr-Gly-Gly-Ile-Ser-Ser (6 residues)

TP-2: Phe-Pro-Asp-Ala-Leu-Lys (6 residues)

30 TP-3: Phe-Asp-Trp-Phe-Lys-Asn-Ala-Asp-Asn-Pro-Ser-Phe-

\* \* \* \* \*

Ser-Phe-Arg (15 residues)

35

[0036] Since the N-terminal amino acid sequence and amino acid sequences obtained by peptide mapping exhibit homology with the amino acid sequence of 43 kDa endoglucanase obtained from Humicola insolens DSM1800 shown in WO 91/17243 (Japanese Patent Laid-Open Publication No. 5-509223), this protein was strongly suggested to be a type of cellulase.

40 [0037] In addition, when the above sequences were compared with the sequences registered in Protein Identification Resource (PIR), R44.0, March 1995 or those registered in SWISS-PROT R31.0, March 1995, although the above sequences exhibited homology with some registered sequences, they were not identical, thus clearly indicating the protein obtained herein to be a novel protein.

## 45 Example A3: Construction of Genome DNA Library

[0038] Isolation of genome DNA was performed in accordance with the method of Horiuchi, et al. (Hiroyuki Horiuchi et al., J. Bacteriol., 170:272-278, 1988).

50 [0039] To begin with, Humicola insolens MN200-1 was cultured in the above (N) medium at 37°C. After culturing for 2 days, the microorganisms were collected by centrifugation (3500 rpm, 10 minutes). The resulting microorganisms were treated with phenol, proteinase K and ribonuclease A followed by polyethylene glycol (PEG) precipitation to obtain genome DNA.

[0040] Then, Humicola insolens genuine DNA was digested by Sau3A I, and after confirming partial digestion within a range of 9-23 kbp by agarose gel electrophoresis, the DNA fragments were collected by ethanol precipitation. The DNA fragments were then ligated to a phage vector and the BamH I arm of an EMBL3 cloning kit (Stratagene) using T4 ligase (Toyobo Co., Ltd.). After subjecting this to ethanol precipitation, the precipitate was dissolved in TE (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) buffer.

55 [0041] The entire amount of ligated mixture was packaged in a lambda head using package components frozen as

described in the method of Hohn, B. (Hohn, B., Methods Enzymol., 68:299-309, 1979) and the Giga-Pack II Packaging Kit (Stratagene), after which the resulting phage was infected into *E. coli* strain LE392. The target gene was then cloned using the 5 x 10<sup>4</sup> phage library resulting from this method.

#### 5 Example A4: Preparation of Long Chain Probe by PCR

[0042] A long probe amplified by PCR was prepared using the total DNA of *Humicola insolens* as a template, and this was used as the DNA probe.

[0043] DNA corresponding to the amino acids indicated with an asterisk on the N-terminal and peptide TP-3 were synthesized for each of the primers. The sequences of the prepared synthetic oligonucleotides were as shown below.

NCE4N1: 5'-GCXGA(CT)GGXAA(AG)TC(AGCT)AC-3' (17mer)

NCE4N2: 5'-GCXGA(CT)GGXAA(AG)AG(CT)AC-3' (17mer)

NCE4C: 5'-CXGC(AG)TT(CT)TT(AG)AACCA(AG)TC-3' (19mer)

15 (X: inosine)

[0044] PCR reactions were conducted under the following conditions. To begin with, 2 pairs of tubes of 1  $\mu$ M each of primers NCE4N1 and NCE4C added to 1  $\mu$ g of *Humicola insolens* genome DNA, and 1  $\mu$ M each of primers NCE4N2 and NCE4C added to 1  $\mu$ g of *Humicola insolens* genome DNA were prepared, and then subjected to thermal denaturation for 5 minutes at 95°C in the presence of dNTP. Later, Taq polymerase (recombinant Taq, Takara Shuzo) was added to amplify by repeating 25 cycles under reaction conditions of 94°C for 1 minute, 45°C for 2 minutes and 72°C for 3 minutes. As a result, approximately 750 bp DNA was amplified only in the case of using primers NCE4N1 and NCE4C. This was used as the probe in following screening.

#### 25 Example A5: Cloning of Cellulase Component NCE4 Gene

##### (1) Screening by Plaque Hybridization

[0045] 100 ng of the approximately 750 bp DNA fragment amplified by PCR was labeled in advance by an ECL Direct DNA/RNA Labeling Detection System (Amersham).

[0046] The phage plaque prepared according to the method described in Example 2 was transferred to High-Bond N + Nylon transfer membrane (Amersham), and after denaturing with 0.4 N sodium hydroxide, it was washed with 5-fold concentrated SSC (15 mM trisodium citrate, 150 mM sodium chloride) and dried to fix the DNA. After pre-hybridizing (42°C) for 1 hour according to the method of the kit, the previously labeled probe was added after which hybridization was performed for 4 hours (42°C). Washing of the label was performed in accordance with method of the above kit. To begin with, washing was repeated twice for 20 minutes at 42°C using 0.5-fold concentrated SSC containing 0.4% SDS and 6 M urea, followed by washing twice for 5 minutes at room temperature using 2-fold concentrated SSC.

[0047] After immersing the Nylon membrane from which the probe had been washed in the detection solution provided for 1 minute, it was photosensitized to Hyper-Film ECL (Amersham) to obtain four positive clones.

##### 40 (2) Preparation of Phage DNA

[0048] *E. coli* LE392 was infected with a phage, the phage particles were collected 8 hours later, and after treating with proteinase K and phenol according to the method of Grossberger (Grossberger, D., Nucleic Acids Res. 15, 6737, 1987), the phage DNA was separated by ethanol precipitation.

##### (3) Subcloning of Target Gene

[0049] Four types of phage DNA were digested by Sal I and applied to agarose gel electrophoresis.

50 [0050] DNA was transferred to a Nylon membrane by the method of Southern (Southern, E.M., J. Mol. Biol., 98: 503-517, 1975), and hybridized using a probe of approximately 750 bp under the same conditions as the plaque hybridization of (1) above followed by detection of the 5.2 kbp DNA fragment containing the target gene. As a result, the four types of phage DNA has Sal I fragments of the same size.

[0051] This 5.2 kbp DNA fragment was separated using the Sephaglass Band Prep Kit (Pharmacia-Biotech) followed by subcloning to the Sal I site of plasmid pUC119 using *E. coli* JN109. The resulting plasmid was designated as pNCE4Sal.

## Example A6: Determination of Base Sequence

## (1) Base Sequence Analysis of Genome DNA

5 [0052] Determination of base sequence was performed in the manner described below.

[0053] The A.L.F. DNA Sequencer II (Pharmacia-Biotech) was used for the base sequence analyzer. Acrylamide carrier able to be acquired as Ready-Mix Gel (Pharmacia-Biotech) or Hydrolink Long Ranger (FMC) was used for the sequencing gel. A.L.F. grade reagents (Pharmacia-Biotech) were used for the various reagents used for gel preparation (N,N,N',N'-tetramethylethylenediamine, urea, ammonium persulphate). The Auto-Read Sequencing Kit (Pharmacia-Biotech) was used for the base sequence decoding reaction. Gel preparation conditions, reaction conditions and migration conditions were set by referring to the details of each instruction manual.

10 [0054] After alkaline-denaturing pNCE4Sal with 10 µg of 2 M sodium hydroxide to use it as a template DNA, it was annealed with the universal primer provided with the Auto-Read Sequencing Kit to perform an elongation reaction. After sequencing the reaction product with the sequencer, a 546 bp base sequence was sequenced. Based on this result, an FITC-labeled sequencing primer in the form of MNEG01 was prepared, reacted with pNCE4Sal followed by further sequencing. The following primer was then prepared based on the result obtained after which the entire length of NCE4 was decoded as a result of proceeding with decoding. The prepared FITC-labeled sequencing primers were as indicated below.

20 MNEG-01: 5'-GTGATGAGGGCTGGCGACAGGCC-3' (23mer)  
MNEG-02: 5'-CTGCCACCTCTATTGCCGGCAGC-3' (23mer)  
MNEG-03: 5'-CCCGACGCCCTCAAGCCCGGCTG-3' (23mer)  
MNEG-04: 5'-GGCTGGAGCGGCTGCACCACTG-3' (23mer)

## 25 (2) Base Sequence Determination

[0055] FITC-labeled sequencing primer DNA in the form of MNEG-05 through MNEG-08 were synthesized based on the results of (1) above. The prepared FITC-labeled sequencing primers were as indicated below.

30 MNEG-05: 5'-GACCTGACGGAAGCTGAAGCTCG-3' (23mer)  
MNEG-06: 5'-AGCAGTGCAGCCGCTGGGAGTCG-3' (23mer)  
MNEG-07: 5'-TGGCAGATGAGGACGTGGTGTG-3' (23mer)  
MNEG-08: 5'-CGCAGCCGGAAGTGGCTCGAAG-3' (23mer)

35 [0056] These primers were reacted with pNCE4Sal using the Auto-Read Sequencing Kit. First of all, 10 µg of plasmid were alkaline-denatured, annealed with each primer, and allowed to react with T7 polymerase. As a result, a 1257 bp base sequence was able to be determined in the Sal I fragment. This sequence was as shown in SEQ ID No. 3.

## Example A7: Determination of Intron

40 [0057] Determination of intron was performed by preparing mRNA from *Humicola insolens* MN200-1, synthesizing cDNA using reverse transcriptase, and comparing this with the genome base sequence to identify the same region.

## (1) Isolation of Whole RNA

45 [0058] *Humicola insolens* MN200-1 was cultured for 2 days in cellulase-inducing medium, and preferable the previously mentioned (N) medium, followed by collection of the microorganisms by centrifugation (3500 rpm, 10 minutes). 2 g of the above microorganisms were washed with sterile water and crushed with a blender (Nihon Seiki, Homogenizer AM-3) while frozen with liquid nitrogen. The crushed microorganisms were then suspended in 10 ml of denaturing solution containing 4 M guanidine thiocyanate (4 M guanidine thiocyanate, 25 mM trisodium citrate, 0.5% sodium N-laurylsarcosinate, 0.1 Mmercaptoethanol). After stirring for several minutes at room temperature, the suspension was neutralized with 1 ml of 2 M sodium acetate (pH 4.5) and additionally stirred after adding 10 ml of TE saturated phenol. 2 ml of chloroform-isoamyl alcohol (24:1) were added and after stirring well, the microorganism components denatured with phenol were removed by centrifugation (3500 rpm, 10 minutes). The upper layer (aqueous layer) was aspirated off and nucleic acids were precipitated with 10 ml of isopropanol. The resulted precipitate was then centrifuged (3500 rpm, 10 minutes) to recover the nucleic acids after which the precipitate was washed with 70% ethanol-water by additional centrifugation.

55 [0059] After dissolving this precipitate in 3.5 ml of TE, 880 µl of 10 M lithium chloride solution were added to the solu-

tion followed by refrigerating for 2 hours at 5°C and centrifugation (12,000 rpm, 10 minutes) to recover the precipitate. This precipitate was washed with 70% ethanol and used as the whole RNA fraction. The recovered amount was 2.7 mg and the yield was 0.14%.

5 (2) Preparation of Poly A Tail+ RNA (= mRNA)

[0060] Preparation of mRNA was performed using a mRNA purification kit (Pharmacia-Biotech).

[0061] To begin with, 1 mg of the whole RNA prepared in (1) above was dissolved in 1 ml of elution buffer followed by thermal denaturation treatment for 10 minutes at 65°C. After rapidly cooling in ice, 0.2 ml of sample buffer were added.  
10 The entire amount of this RNA solution was charged into an oligo(dT) cellulose column, and after washing the column three times with high-salt buffer and 3 times with low-salt buffer, eluted with elution buffer heated to 65°C. This column procedure was repeated twice and the resulting product was used as the mRNA fraction. The amount recovered was 19.2 µg the yield was 2%.

15 (3) Synthesis of cDNA

[0062] Synthesis of cDNA was performed using a time-saver cDNA synthesis kit (Pharmacia-Biotech).

[0063] 5 µg of mRNA were dissolved in 20 µg of sample buffer. After thermal denaturing for 10 minutes at 65°C, dithiothreitol solution and oligo(dT) primer were both added to first strand synthesis mix and allowed to react for 1 hour at 37°C. Then, this entire amount was added to second strand mix and allowed to react for 30 minutes at 12°C and then for 1 hour at 22°C, and the resulting product was used as cDNA.

(4) PCR Amplification of Cellulase NCE4 cDNA

25 [0064] Using 1 µg of the synthesized cDNA as a template, only the target cDNA was amplified by PCR. Oligonucleotide primers having the sequences indicated below were prepared for use as primers of the N- and C-terminals.

NCE4-CN: 5'-ATGCGTTCCTCCCTCTCCTCCGCTCCGCC-3' (30mer)

NCE4-CC: 5'-TACAGGCACTGATGGTACCAGTCATTAATC-3' (30mer)

30 [0065] The PCR reaction was carried out under the following conditions. 1 µM each of primer was added to 1 µg of Humicola insolens cDNA followed by thermal denaturation for 10 minutes at 94°C in the presence of dNTP. Then, Taq polymerase (recombinant Tag, Takara Shuzo) was added to amplify by repeating 30 cycles under reaction conditions of 94°C for 1 minute, 50°C for 2 minutes and 72°C for 3 minutes. As a result of performing agarose gel electrophoresis,  
35 the amplified fragment was determined to be 0.9 kbp in size. This was then concentrated by ethanol precipitation and cloned using a pT7 blue-T vector kit (Novagen). This plasmid was designated as pCNCE4.

(5) Analysis of cDNA Base Sequence

40 [0066] The sequencing reaction was performed using an Auto-Read Sequencing Kit in the same manner as previously described.

[0067] Plasmid pCNCE4 was alkaline-denatured with 2M sodium hydroxide and precipitated with ethanol. This single-strand plasmid was used as a template and allowed to react with T7 polymerase. The sequence was determined by performing reactions using the above synthetic primers MNEG01, MNEG02, MNEG03, MNEG04, MNEG05, MNEG06,  
45 MNEG07 and MNEG08 as well as the universal primer and reverse primer provided with the kit.

[0068] As a result, a single, 56 bp intron was present. The non-translated initiation sequence, its stop sequence, and the regulatory sequence within the intron in the sequence of SEQ ID No. 3 are as shown below (the numbers indicate the sequence position numbers of SEQ ID No. 3).

50 Intron: 453-458, 506-508, 491-497

Example A8: Evaluation of the Nap Removing Activity of NCE4 on Cellulose-Containing Fibers

[0069] Pre-dyca Tencel fabric (Courtauls) was napped using a large washer adding surfactant and a rubber ball. Tencel in which nap were formed as a result of this treatment was removed of nap by cellulase treatment under the following conditions followed by calculation of the protein concentration of cellulase required to completely remove the nap.

Test machine: 20 kg washer (Sanyo, Fully-Automated Washing Machine SCW5101)



Bath ratio: 1:20  
 Heating: 55°C  
 Time: 60 minutes  
 pH: 7 (10 mM phosphate buffer)

[0070] In addition to cellulase enzyme liquid, rubber balls were added in an amount roughly twice the weight of the cloth to the treatment liquid.

Table 1

	Amt. of Protein Required to Remove Nap from Tencel*
NCE4	5 mg/liter
Crude Cellulase Preparation	500 mg/liter

\*: The amount of protein was quantified by using a protein assay kit (Biorad) and bovine serum albumin for the standard.

[0071] Based on the results, it was determined that NCE4 is able to demonstrate an equal degree of Tencel nap removal at a protein concentration 1/100 that of crude cellulase preparation.

Example A9: Evaluation of Decoloring Activity of NCE4 on Denim-Dyed, Cellulose-Containing Fibers

[0072] Destarched 12-ounce blue jeans were decolorized under the conditions described below.

Test machine: 20 kg washer (Sanyo, Fully-Automatic Washing Machine SCW5101)  
 Bath ratio: 1:50  
 Heating: 55°C  
 Time: 60 minutes  
 pH: 7 (10 mM phosphate buffer)

[0073] In addition to cellulase enzyme liquid, rubber balls were added in an amount roughly twice the weight of the cloth to the treatment liquid.

[0074] Decoloring was determined by measuring the L value (lightness) of an Lab indication system using a color difference type of Color Analyzer Topscan Model TC-1800MK2 (Tokyo Denshoku). Representing the increase in L value relative to a control (increase in lightness) as the  $\Delta L$  value, the  $\Delta L$  values of 5 points for each test area evaluated for decoloring were measured ( $n = 5$ ), and the mean value of 3 points after discarding the maximum and minimum values was used. The protein concentration of cellulase required to decolor to a level of  $\Delta L = 7$  was calculated.

Table 2

	Amt. of Protein Required to Discolor Blue Jeans
NCE4	1.8 mg/liter
Crude cellulase preparation	45.0 mg/liter

[0075] Based on these results, NCE4 was determined to be able to demonstrate an equal degree of discoloring of denim-dyed blue jeans at a protein concentration 1/25 that of crude cellulase preparation.

Example A10: Evaluation of Weight Reduction Activity of NCE4 on Cellulose-Containing Fibers

[0076] Enzyme treatment was performed on Cupra, Recycled cellulose fibers (Asahi Chemical Industry, 15 cm long x 10 cm wide) for which absolute dry weight was measured in advance under the conditions indicated below.

## EP 0 959 128 A1

Test machine: Washing resistance tester L-12 (Daiei Scientific Instruments)

Bath ratio: 1:50

Heating: 55°C

Time: 60 minutes

pH: 7 (40 mM phosphate buffer)

[0077] In addition to cellulase enzyme liquid, stainless steel balls provided (Daiei Scientific Instruments) were added to the treatment liquid. Following enzyme treatment, the fibers were dried, the absolute dry weight of the Cupra fibers were measured, and the weight reduction ratio relative to before enzyme treatment was measured. The protein concentration of cellulase required to produce a weight reduction ratio of 8% was calculated.

Table 3

	Amt. of Protein Required for Cupra Weight Reduction of 8%
NCE4	100 mg/liter
Crude cellulase preparation	500 mg/liter

[0078] Based on these results, NCE4 was determined to be able to demonstrate an equal degree of weight reduction processing on Cupra fibers at a protein concentration 1/5 that of crude cellulase preparation.

### Example B1: Construction of Plasmid pMKD01

#### (1) Construction of Plasmid pUC118BN

[0079] After cleaving 1 µg of pUC118 DNA with BamH I, the restriction enzyme was inactivated by treatment with phenol. This was then subjected to ethanol precipitation and dissolved in a small amount of TE (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) buffer. The ends of this DNA were blunted with a DNA branching kit (Takara Shuzo). This was further ligated with a DNA ligation kit (Takara Shuzo) and allowed to self-close. The resulting ligated mixture was transformed into E. coli component cells JM109 (Takara Shuzo). Those transformants able to grow on LB agar medium (1% polypeptone, 0.5% yeast extract, 1% NaCl, 1.5% agar) containing 100 µg/ml of ampicillin, 1 mM IPTG and 0.004% of X-gal that also demonstrated white colonies were selected, and additionally cultured overnight at 37°C in LB medium (1% polypeptone, 0.5% yeast extract, 1% NaCl) containing 100 µg/ml of ampicillin. Plasmid DNA was recovered from the resulting culture liquid using alkaline-SDS. This plasmid DNA was cleaved with BamH I and applied to 0.8% agarose gel electrophoresis followed by selection of plasmid DNA in which the BamH I site of pUC118 DNA was destroyed. This plasmid DNA was designated as pUC118BN.

#### (2) Construction of Plasmid pUC118BSN

[0080] 1 µg of pUC118BN DNA was digested with Sph I and plasmid DNA was obtained in which the Sph I site of pUC118BN was destroyed according to the same method as described above. This plasmid DNA was designated as pUC118BSN.

#### (3) Construction of Plasmid pM21

##### A) Isolation of Cellulase NCE2 Gene

[0081] The cellulase NCE2 gene obtained from Humicola insolens according to the method described in Japanese Patent Laid-Open Publication No. 8-126492, and a Pst I-Xba I fragment of a total length 3.4 Kbp having a 1.4 Kb DNA sequence upstream and a 0.5 Kb DNA sequence downstream as its promoter and terminator regions, were ligated to the Pst I-Xba I site of the previous pUC118BSN. The resulting plasmid DNA was designated as pUC118BSN-PX.

##### B) Site-directed mutagenesis reaction of Plasmid pUC118BSN-PX

[0082] BamH I sites were introduced by site-specific mutation as described below downstream of the N-terminal and

immediately downstream from the stop codon of NCE2 gene. *E. coli* JM109 strain was transformed by plasmid pUC118BSN-PX, and after additionally infecting with helper phage M13K07, culturing was performed for 16-20 hours at 37°C in 30 ml of 2xYT liquid medium (1.6% bactotryptone, 0.8% yeast extract, 0.5% NaCl) containing ampicillin and kanamycin at concentrations of 150 µg/ml and 70 µg/ml, respectively. Single-strand DNA of M13 (ssDNA) was recovered from the culture supernatant. Site-directed mutagenesis reaction was performed using this ssDNA and two types of synthetic oligonucleotides with the Sculpture In Vitro Mutagenesis System (Amersham). The sequences of the prepared synthetic oligonucleotide primers are as shown below.

MNC-02 5'-GAGCGCCAGAACTGTGGATCCACTTGGTGAGCAATG-3' (36mer)

MNC-03 5'-TCCGCCGTTCTGAGCGGATCCAGCGCTTTGGCGCG-3' (35mer)

[0083] The mixture of site-directed mutated DNA was introduced into *E. coli* TG1 and the resulting transformant was cultured in LB medium (1% polypeptone, 0.5% yeast extract, 1% NaCl) containing 100 µg/ml of ampicillin, followed by recovery of the plasmid DNA. This plasmid DNA was digested with BamH I and applied to 0.8% agarose gel electrophoresis followed by selection of plasmid DNA in which BamH I sites were introduced at two locations in plasmid pUC118BSN-PX. This plasmid DNA was designated as pM21.

#### (4) Isolation of The Cellulase NCE3 Gene

[0084] The cellobiohydrolase gene (NCE3) originating in *Humicola insolens* was isolated by PCR based on the sequence of known cellobiohydrolase gene originating in *Humicola grisea* (de Oliveira Alzevedo. M. et al., J. General Microbiol., 136: 2569-2576, 1990).

##### A) Isolation of Genome DNA

[0085] Genome DNA of *Humicola insolens* MN 200-1 was obtained by the method of previously described Example A3.

##### B) Amplification of The Cellulase NCE3 Gene by PCR

[0086] The NCE3 gene of *Humicola insolens* was isolated by PCR based on the sequence of cellobiohydrolase gene originating in *Humicola grisea*. Each primer was designed in advance in a form that contained a BamH I site so that the PCR product containing NCE3 is able to be ligated in line with the frame to the BamH I site of plasmid pM21. Synthetic oligonucleotides having the sequences shown below were prepared for use as primers.

MKA-05: 5'-GCCGCCAGCAGGCGGGATCCCTCACCACCGAGAGG-3' (36mer)

MKA-06: 5'-TGATCGTCGAGTCAGGGATCCAGAAITTTACAGGCAC-3' (36mer)

[0087] The PCR reaction was carried out according to the following method in accordance with the LA PCR Kit Ver. 2 (Takara Shuzo). To begin with, 1 µM of each primer, 400 µM dNTP and 2.5 U of LA Taq polymerase were added to 1 µg of the *Humicola insolens* genome DNA obtained by the method described above followed by amplification by repeating 30 cycles under reaction conditions of 94°C for 1 minute, 55°C for 2 minutes and 72°C for 3 minutes. As a result of 0.8% agarose gel electrophoresis, amplification of 1.6 Kbp DNA was confirmed. This 1.6 Kbp DNA fragment was recovered using a Sephaglass Band Prep Kit (Pharmacia-Biotech), and then ligated to pT7 blue-T vector kit (Novagen). This plasmid DNA was designated as pK21.

#### (5) Construction of Plasmid pKM04

[0088] Plasmid pK21 DNA was digested with BamH I and a 1.6 Kbp DNA segment was recovered. Then, plasmid pM21 DNA was digested with BamH I after which restriction enzyme was inactivated by treating for 10 minutes at 70°C. This was dephosphated with calf alkaline phosphatase (Takara Shuzo) and separated by 0.8% agarose gel electrophoresis followed by recovery of a 5.2 Kbp DNA fragment. The 1.6 Kbp DNA fragment originating in pK21 and the 5.2 Kbp DNA fragment originating in pM21 were ligated to obtain plasmid pKM04.

#### (6) Construction of Plasmid pMKD01

[0089] To begin with, a gene was prepared that allowed expression of the destomycin-resistant gene described in Japanese Patent Laid-Open Publication No. 59-175889 in *Humicola insolens* using the promoter and terminator of the trp

C gene originating in *Aspergillus nidulans* obtained by known methods (Mullaney, E.J. et al., Mol. Gen. Genet., 199: 37-45, 1985). This gene was introduced into the Xba I site of plasmid pKM04 to prepare plasmid pMKD01.

#### Example B2: Transformation of *Humicola insolens* by Plasmid pMKD01

##### (1) Preparation of Highly Concentrated Purified Standard of Plasmid pMKD01

[0090] Highly concentrated purified standard of pMKD01 was first prepared in order to introduce plasmid pMKD01 into *Humicola insolens*. After introducing pMKD01 into *E. coli* JM109, it was cultured overnight at 37°C in 100 ml of LB medium containing 100 µg/ml of ampicillin. The resulting culture liquid was purified using a Flexiprep Kit (Pharmacia-Biotech) to obtain 1 µg/µl of pMKD01 plasmid DNA.

##### (2) Transformation of *Humicola insolens*

[0091] *Humicola insolens* MN200-1 was cultured at 37°C in (S) medium after which the microorganisms were collected by centrifugation for 10 minutes at 3000 rpm 24 hours later. Here, the composition of the (S) medium consisted of adding glucose (3.0%) but omitting Avicel from the previously mentioned (N) medium. The resulting microorganisms were washed with 0.5 M sucrose and suspended in 10 ml of protoplasmic enzyme solution filtered with a 0.45 µm filter (5 mg/ml Novozyme 234 (NLI), 5 mg/ml Cellulase Onozuka R-10 (Yakult), 0.5 M sucrose). The mycelia were converted into protoplasts by shaking for 60 to 90 minutes at 30°C. After filtering this suspension, the protoplasts were recovered by centrifugation for 10 minutes at 2500 rpm and then washed with SUTC buffer (0.5 M sucrose, 10mM calcium chloride, 10mM Tris-HCl (pH 7.5)).

[0092] Protoplasts prepared in the above manner were suspended in 1 ml of SUTC buffer followed by the addition of 10 µg of DNA (TE) solution (10 µl) to 100 µl of this suspension and allowing to stand undisturbed for 5 minutes in ice. Then, 400 µl of PEG solution (60% PEG4000, 10 mM calcium chloride, 10 mM Tris-HCl (pH 7.5)) were added, and after allowing to stand undisturbed for 20 minutes in ice, 10 ml of SUTC buffer were added followed by centrifugation for 10 minutes at 2500 rpm. After suspending the collected protoplasts in 1 ml of SUTC buffer, they were again centrifuged for 5 minutes at 4000 rpm and finally suspended in 100 µl of SUTC buffer.

[0093] Protoplasts subjected to the above treatment were layered with YMG soft agar over YMG medium containing 200 µg/ml of hygromycin B (1% glucose, 0.4% yeast extract, 0.2% malt extract, 1% agar (pH 6.8)), and after culturing for 5 days at 37°C, the formed colonies were used as transformants.

##### (3) Culturing of pMKD01 Transformed Strains and Evaluation by SDS-PAGE

[0094] Plasmid pMKD01 was introduced into *Humicola insolens* MN200-1 as previously described followed by selection of 50 strains exhibiting resistance to hygromycin. These were cultured for 5 days at 37°C in (N) medium. When the resulting culture supernatant was analyzed by SDS-PAGE, in 5 clones of pMKD01 transformed strains, the protein band estimated to be NCE3 was increased 3-4 times higher than the parent strain.

##### (4) Identification of N-Terminal Amino Acid Residues of Recombinant NCE3

[0095] In order to confirm that the protein band expressed in large amounts originated in the NCE3 gene based on the results of SDS-PAGE, the N-terminal amino acid sequence of this protein was determined. To begin with, column chromatography was performed on the culture supernatants obtained from the parent strain and NCE3 high-expression strains using an FPLC system in accordance with the method of the previously described Example A2, followed by comparison of the major peaks. The peak that was particularly increased in the NCE3 high-expression strains was removed and lyophilized. This was dissolved in a small amount of water followed by electrophoresis using 8% Gel SDS-PAGE Mini (Difco). The protein was then electrically transferred to a PVDF membrane according to the method of Example A2 above, and after staining with Coomassie brilliant blue R-250, the membrane was decolorized and washed with water. A portion in which a protein having a molecular weight of 66 KD was blotted was cut out from this membrane, and the modified N-terminal residues were removed from the blotted protein according to the method of Podell, D.N. et al. (Podell, D.N., et al., Biochem. Biophys. Res. Commun., 81: 176, 1978). First of all, the target protein was cut out, and after incubating for 30 minutes at 37°C in a small amount of 0.5% polyvinylpyrrolidone (molecular weight: 40,000, Sigma)/100 mM acetic acid solution, was washed well with water. Then, the modified N-terminal residues were removed by Pfu pyroglutamate aminopeptidase (Takara Shuzo), washed with water and air dried. This was then applied to a Model 492 Protein Sequencer to determine 15 residues of the N-terminal amino acid sequence. The resulting sequence is shown below.

## EP 0 959 128 A1

N-terminal amino acid sequence: Asn-Cys-Gly-Ser-Leu-Thr-Thr-Glu-Arg-His-Pro-Ser-Leu-Ser-Trp (15 residues)

[0096] This N-terminal amino acid sequence coincided with the cellulase NCE2 and NCE3 fused protein amino acid sequence estimated from the base sequence of plasmid pMKD01.

### (5) FPLC Evaluation of pMKD01 Transformed Strains

[0097] As was previously described, column chromatography was performed using an FPLC system to further quantify the culture supernatants of the 5 clones confirmed to express large amounts of NCE3 by SDS-PAGE. The conditions were the same as in (4) above. The NCE3 peak was removed and lyophilized, and the weight was measured and the productivity was compared between high expression strains and the parent strain. Those results were as shown in the table below.

Table 4

	NCE3 Productivity*
Humicola insolens MN200-1 (parent strain)	0.46 g
Humicola insolens pMKD01	1.8 g

\* Productivity refers to the amount produced per one liter of culture liquid.

### Example B3: Construction of Plasmid pEGD01

[0098] Plasmid pMKD01 was digested with BamH I and restriction enzyme was inactivated by heat treatment at 70°C after which it was dephosphated to recover an 8.2 Kbp DNA fragment.

[0099] Then, the NCE4 gene was amplified by PCR based on the sequence of the NCE4 gene originating in Humicola insolens obtained in the above Examples A1 through A7. This PCR product containing NCE4 was designed in a form in which each primer contained a BamH I site in advance so that it could be ligated while aligning the frame with the 8.2 Kbp BamH I fragment of the above plasmid pMKD01. Synthetic oligonucleotides having the sequences shown below were prepared for use as primers.

NCE4-N: 5'-CCGGTGTGGCCGGATCCGCTGATGGCAAG-3' (30mer)

NCE4-C: 5'-TAAGGCCCTCAAGGATCCCTGCGTCTACAG-3' (30mer)

[0100] The PCR reaction was carried out in the following manner. 1  $\mu$ M of each primer, 400  $\mu$ M dNTP and 2.5 U of Pfu DNA polymerase (Stratagene) were added to 1  $\mu$ g of Humicola insolens genome DNA, and a 0.8 Kbp DNA fragment was amplified by repeating 25 cycles of reaction conditions consisting of 94°C for 1 minute, 55°C for 2 minutes and 72°C for 3 minutes. This 0.8 Kbp DNA fragment was recovered and ligated to the above 8.2 Kbp BamH I fragment of pMKD01. This plasmid DNA was designated as pEGD01.

### Example B4: Expression of Plasmid pEGD01

#### (1) Transformation of Humicola Insolens by Plasmid pEGD01

[0101] Transformation of Humicola insolens MN200-1 by plasmid pEGD01 was performed in accordance with the method of Example B2. To begin with, a highly concentrated and purified standard of pEGD01 was prepared to obtain 1  $\mu$ g/ $\mu$ l of pEGD01 plasmid DNA. Humicola insolens MN200-1 was transformed using 10  $\mu$ l of this pEGD01 solution followed by selection of 50 transformant strains exhibiting resistance to hygromycin. These were cultured for 5 days at 37°C in (N) medium. When the resulting culture supernatant was analyzed by SDS-PAGE, the protein band estimated to be NCE4 was increased 10-16 times more than in the parent strain in 10 clones of the strains transformed by pEGD01.

#### (2) Identification of N-Terminal Amino Acid Residues of Recombinant NCE4

[0102] In order to confirm that protein band expressed in large amounts originated in the NCE4 gene based on the results of SDS-PAGE, the N-terminal amino acid sequence of this protein was determined. To begin with, column chro-

matography was performed on the culture supernatants obtained from the parent strain and NCE4 high-expression strains using an FPLC system followed by comparison of the major peaks. The conditions were the same as in the above Example B2. The peak that was particularly increased in the NCE4 high-expression strains was removed and lyophilized. This was dissolved in a small amount of water. After removing the modified N-terminal residues according to the method of Example B2, the N-terminal amino acid sequence was determined using the above protein sequencer. As a result, two types of N-terminal amino acid sequences were obtained in a ratio of about 7:3. The resulting sequences were as shown below. Then, the N-terminal amino acid sequence was determined using the same protein sequencer without removing the modified N-terminal. As a result, only amino acid sequence 1 was obtained as shown below.

N-terminal amino acid sequence 1: Val-Val-Glu-Glu-Arg-Gln-Asn-Cys-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp (20 residues)

N-terminal amino acid sequence 2: Asn-(Cys)-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp-(Cys)-(Cys)-Lys-Pro-Ser-(Cys) (20 residues)

[0103] These N-terminal amino acid sequences coincided with the cellulase NCE2 and NCE4 fused protein amino acid sequence estimated from the base sequence of plasmid pEGD01. Since two types of N-terminal amino acid sequences were obtained, it was clearly shown that when the signal sequence of this fused protein is cleaved, it is processed at a plurality of sites.

### (3) FPLC Evaluation of pEGD01 Transformed Strains

[0104] As was previously described, column chromatography was performed using an FPLC system to further quantify the culture supernatants of the 5 clones confirmed to express large amounts of NCE4 by SDS-PAGE, and the NCE4 peak was removed. This was then lyophilized, and the weight thereof was measured and the productivity was compared between high expression strains and the parent strain. Those results were as shown in the table below.

Table 5

	NCE4 Productivity*
Humicola insolens MN200-1 (parent strain)	0.28 g
Humicola insolens pMKEG1	4.5 g

\* Productivity refers to the amount produced per one liter of culture liquid.

### Example B5: Construction of Plasmid pIED02

#### (1) Construction of Plasmid pID01

[0105] Plasmid pEGD01 was digested with Hind III and BamH I and a 7.2 Kbp DNA fragment was recovered.

[0106] Then, the portion of DNA that codes for the promoter and signal sequences of NCE1 gene was amplified by PCR based on the sequence of NCE1 gene originating in the Humicola insolens obtained with the method described in Japanese Patent Laid-Open Publication No. 8- 56663. This PCR product that contains NCE1 promoter and signal sequence was designed in a form in which each primer contains an Hind III and BamH I site in advance so that it can be ligated to the 7.2 Kbp Hind III to BamH I fragment of the above plasmid pEGD01. Synthetic oligonucleotides having the sequences shown below were prepared for use as primers.

PNCE1-N: 5'-GTCATGAAGCTTCATTAAGGTACGTATGCAAC-3' (32mer)

PNCE1-C: 5'-GGTGATGGATCCGGCCTGCTGGCAGCGACGC-3' (32mer)

[0107] The PCR reaction was carried out in the same method as Example B3. 1  $\mu$ M of each primer, 400  $\mu$ M dNTP and 2.5 U of Pfu DNA polymerase were added to 1  $\mu$ g of Humicola insolens genome DNA, and a 1.5 Kbp DNA fragment was amplified by repeating 23 cycles of reaction conditions consisting of 94°C for 1 minute, 55°C for 2 minutes and 72°C for 4 minutes. This 1.5 Kbp DNA fragment was recovered by digesting the PCR product with Hind III and BamH I. It was then ligated to the above 7.2 Kbp Hind III to BamH I fragment of pEGD01. This plasmid DNA was designated as pID01.

## (2) Construction of Plasmid pIED02

[0108] Plasmid pID01 was digested with BamH I and restriction enzyme was inactivated by heat treatment at 70°C after which it was dephosphated to recover an 8.6 Kbp DNA fragment. Then, after digesting plasmid pEGD01 with BamH I, a 0.8 Kbp DNA fragment was recovered that contained NCE4 gene. The two fragments were ligated to obtain plasmid pIED02.

## Example B6: Expression of Plasmid pIED02

(1) Transformation of *Humicola insolens* by Plasmid pIED02

[0109] Transformation of *Humicola insolens* MN200-1 by plasmid pIED02 was performed in accordance with the method of Example B2. To begin with, a highly concentrated and purified standard of pIED02 was prepared to obtain 1 µg/µl of pIED02 plasmid DNA. *Humicola insolens* MN200-1 was transformed using 10 µl of this pIED02 solution followed by selection of 50 transformant strains exhibiting resistance to hygromycin. These were cultured for 5 days at 37°C in (N) medium. When the resulting culture supernatant was analyzed by SDS-PAGE, the protein band estimated to be NCE4 was increased 5-10 times more than in the parent strain in 5 clones of the strains transformed by pIED02.

## (2) Identification of N-Terminal Amino Acid Residues of Recombinant NCE4

[0110] In order to confirm that protein band expressed in large amounts originated in NCE4 gene based on the results of SDS-PAGE, the N-terminal amino acid sequence of this protein was determined. To begin with, according to the method of Example B2 column chromatography was performed on the culture supernatants obtained from the parent strain and the NCE4 high-expression strains, after which the NCE4 peak was removed and lyophilized. This was dissolved in a small amount of water. After removing the modified N-terminal residues according to the method of Example B2, 15 residues of the N-terminal amino acid sequence were determined using the above protein sequencer. The resulting sequence was as shown below.

N-terminal amino acid sequence: Gln-Ala-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp-(Cys) (15 residues)

[0111] This N-terminal, amino acid sequence coincided with the cellulase NCE1 and NCE4 fused protein amino acid sequence estimated from the base sequence of plasmid pIED02.

## (3) FPLC Evaluation of pIED02 Transformed Strains

[0112] As was previously described, column chromatography was performed using an FPLC system to further quantify the culture supernatants of the 5 clones confirmed to express large amounts of NCE4 by SDS-PAGE, and the NCE4 peak was removed. This was then lyophilized, and the weight was measured and the productivity was compared between high expression strains and the parent strain. Those results were as shown in the table below.

Table 6

	NCE4 Productivity*
<i>Humicola insolens</i> MN200-1 (parent strain)	0.28 g
<i>Humicola insolens</i> pIED02	2.9 g

\* Productivity refers to the amount produced per one liter of culture liquid.

## Sequence Listing

SEQ ID No.: 1  
 Length: 305  
 Type: Amino acid  
 Topology: linear  
 Molecule type: Protein  
 Sequence description:  
 Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro  
           -20                          -15                          -10  
 Val Leu Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys  
           -5                          1                          5                          10  
 Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro  
                           15                          20                          25  
 Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala  
                           30                          35                          40  
 Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln  
 25          45                          50                          55  
 Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr  
 60                          65                          70                          75  
 Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu  
 30                          80                          85                          90  
 Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln  
                           95                          100                          105  
 Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn  
 35          110                          115                          120  
 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe  
 40          125                          130                          135  
 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu  
 140                          145                          150                          155  
 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe  
 45                          160                          165                          170  
 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val  
                           175                          180                          185  
 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp  
 50          190                          195                          200



Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser  
 205 210 215  
 5 Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Ser Thr Ser Thr Thr  
 220 225 230 235  
 Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu  
 240 245 250  
 10 Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys  
 255 260 265  
 Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys  
 270 275 280  
 15 Leu

SEQ ID No.: 2  
 Length: 1257  
 20 Type: Nucleic acid  
 Strandness: double  
 Topology: linear  
 Molecule type: Genomic DNA  
 25 Original source:  
 Organism: *Humicola insolens*  
 Feature:  
 30 Feature Key: Intron  
 Location: 453..509  
 Method of determining feature: E  
 Sequence description:  
 35 AATGACGGGG CAACCTCCCG CCCGGGCCCA ACTCTGGGT TTGTTTGAC AGGCCGTCTG 60  
 TCTCTTGGCT CCTCTTACTA CGCCTGCCTG GACCCTACGT CTCAACTCCG ATTCAAG 117  
 ATG CGT TCC TCC CCT CTC CTC CGC TCC GCC GTT GTG GCC GCC CTG CCG 165  
 Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro  
 40 -20 -15 -10  
 GTG TTG GCC CTT GCC GCT GAT GGC AAG TCC ACC CGC TAC TGG GAC TGC 213  
 Val Leu Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys  
 -5 1 5 10  
 45 TGC AAG CCT TCG TGC GGC TGG GCC AAG AAG GCT CCC GTG AAC CAG CCT 261  
 Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro  
 15 20 25

50  
 55

GTC TTC TCC TGC AAC GCC AAC TTC CAG CGT CTC ACT GAC TTC GAC GCC 309  
 Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala  
 5 30 35 40  
 AAG TCC GGC TGC GAG CCG GGC GGT GTC GCC TAC TCG TGC GCC GAC CAG 357  
 Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln  
 45 50 55  
 10 ACC CCA TGG GCT GTG AAC GAC GAC TTC GCG TTC GGT TTT GCT GCC ACC 405  
 Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr  
 60 65 70 75  
 15 TCT ATT GCC GGC AGC AAT GAG GCG GGC TGG TGC TGC GCC TGC TAC GA 452  
 Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Gl  
 80 85 90  
 GTAAGCTTTG GTCGCGTGTG TAACACTGTG CAGGCATAGC ACTAACCACC TCCCAG G 509  
 20 u  
 CTC ACC TTC ACA TCC GGT CCT GTT GCT GGC AAG AAG ATG GTC GTC CAG 557  
 Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln  
 95 100 105  
 25 TCC ACC AGC ACT GGC GGT GAT CTT GGC AGC AAC CAC TTC GAT CTC AAC 605  
 Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn  
 110 115 120  
 30 ATC CCC GGC GGC GGC GTC GGC ATC TTC GAC GGA TGC ACT CCC CAG TTC 653  
 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe  
 125 130 135  
 GGC GGT CTG CCC GGC CAG CGC TAC GGC GGC ATC TCG TCC CGC AAC GAG 701  
 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu  
 35 140 145 150  
 TGC GAT CGG TTC CCC GAC GCC CTC AAG CCC GGC TGC TAC TGG CGC TTC 749  
 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe  
 160 165 170  
 40 GAC TGG TTC AAG AAC GCC GAC AAC CCG AGC TTC AGC TTC CGT CAG GTC 797  
 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val  
 175 180 185  
 45 CAA TGC CCA GCC GAG CTC GTC GCT CGC ACC GGA TGC CGC CGC AAC GAC 845  
 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp  
 190 195 200

50

55

5 GAC GGC AAC TTC CCT GCC GTC CAG ATC CCC TCC AGC AGC ACC AGC TCT 893  
 Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser  
 205 210 215  
 CCG GTC GGC CAG CCT ACC AGT ACC AGC ACC ACC TCC ACC TCC ACC ACC 941  
 Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Ser Thr Ser Thr Thr  
 220 225 230 235  
 10 TCG AGC CCG CCC GTC CAG CCT ACG ACT CCC AGC GGC TGC ACT GCT GAG 989  
 Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu  
 240 245 250  
 15 AGG TGG GCT CAG TGC GGC GGC AAT GGC TGG AGC GGC TGC ACC ACC TGC 1037  
 Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys  
 255 260 265  
 GTC GCT GGC AGC ACC TGC ACG AAG ATT AAT GAC TGG TAC CAT CAG TGC 1085  
 20 Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys  
 270 275 280  
 CTG TAA ACGCAGGGCA GCCTGAGAAC CTTACTGGTT GCGCAACGAA ATGACACTCC 1141  
 Leu  
 25 CAATCACTGT ATTAGTTCTT GTACATAATT TCGTCATCCC TCCAGGGATT GTCACATATA 1201  
 TGCAATGATG AATACTGAAC ACAAACCTGG CCGCTTGAAC TGGCCGAAGG AATGCC 1257

30

### Claims

1. A protein or modified protein having a portion of an amino acid sequence shown in SEQ ID No. 1 or a sequence from position 1 to position 284 of the amino acid sequence shown in SEQ ID No. 1.
2. A protein or modified protein according to claim 1 further comprising a portion or all of the amino acid sequence from position -21 through -1 shown in SEQ ID No. 1 at N-terminal thereof.
3. A protein or modified protein according to claims 1 or 2 having endoglucanase activity.
4. A DNA sequence coding for the protein or modified protein according to any one of claims 1 to 3.
5. A DNA sequence according to claim 4, wherein the DNA sequence has a portion or all of the base sequence shown in SEQ ID No. 2.
6. A DNA sequence according to claim 5 having a base sequence from position 118 to position 1088 of the base sequence shown in SEQ ID No. 2.
7. A vector comprising the DNA sequence according to any one of claims 4 to 6.
8. A host cell transformed by the vector according to claim 7.
9. A method for preparing a protein or modified protein according to any one of claims 1 to 3, comprising the steps of culturing the host cell according to claim 8 to obtain a culture, and collecting the protein or modified protein according to any one of claims 1 to 3 from the culture.
10. A cellulase preparation comprising a protein or modified protein according to any one of claims 1 to 3.

11. A method for removing nap of cellulose-containing fibers, the method comprising contacting the protein or modified protein according to any one of claims 1 to 3 or the cellulase preparation according to claim 7 with cellulose-containing fibers.
- 5 12. A method for decoloring denim-dyed, cellulose-containing fibers, the method comprising contacting the protein or modified protein according to any one of claims 1 to 3 or the cellulase preparation according to claim 7 with denim-dyed, cellulose-containing fibers.
- 10 13. A method for reducing cellulose-containing fibers, the method comprising contacting the protein or modified protein according to any one of claims 1 to 3 or the cellulase preparation according to claim 7 with cellulose-containing fibers.
- 15 14. Use of the protein or modified protein according to any of claims 1 to 3 or the cellulase preparation according to claim 7 for removing nap of cellulose-containing fibers.
- 16 15. Use of the protein or modified protein according to any one of claims 1 to 3 or the cellulase preparation according to claim 7 for decoloring denim-dyed, cellulose-containing fibers.
- 20 16. Use of the protein or modified protein according to any one of claims 1 to 3 or the cellulase preparation according to claim 7 for reducing cellulose-containing fibers.

25

30

35

40

45

50

55

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP97/02561

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
Int. C16 C12N9/42, C12N15/56		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
Int. C16 C12N9/42, C12N15/56		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
CAS, BIOSIS PREVIEWS, WPI, GENETYX MAC/DB		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO, 96/00787, A1 (Novo Nordisk Biotech, Inc.), January 11, 1996 (11. 01. 96) & EP, 777737, A1 & FI, 9605200, A	1 - 16
X	WO, 96/17994, A1 (Novo Nordisk), June 13, 1996 (13. 06. 96) & AU, 9539791, A	1 - 16
Y	Charlotte SCHOU et al., "Stereochemistry, specificity and kinetics of the hydrolysis of reduced cellodextrins by nine cellulases" Eur. J. Biochem., 1993, Vol. 217, p. 947-953	1 - 16
Y	JP, 8-501692, A (Novo Nordisk A/S), February 27, 1996 (27. 02. 96) & WO, 94/07998, A1 & EP, 663950, A1 & FI, 9501629, A	1 - 16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
September 24, 1997 (24. 09. 97)		October 7, 1997 (07. 10. 97)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.		Authorized officer  Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)